STAT

AN ASSESSMENT OF THE UTILITY OF THE ARPA NETWORK OF COMPUTERS FOR THE INTERNATIONAL SECURITY AFFAIRS ANALYST

Working Paper

1 September 1973

Principal Investigator:
Robert A. Young

Sponsored by
Defense Advanced Research Projects Agency

Contract No.: DAHC15-71-C-0201
ARPA Order No.: 2323
Program Code No.: 2D166

Approved For Release 2004/09/23 : CIA-RDP79M00096A000500010013-4

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Cont'd)

Approved For Release 2004/09/23 : CIA-RDP79M00096A000500010013-4

The ARPA Network is one of the most important technologically innovative ways of using computers since their wide spread introduction in the 1950's. It permits many different users in different fields to expand the scope of their computing tasks far beyond what could be done with any single computer system. The ARPANET has provided practical assistance to some users with particular research problems. It is now appropriate, however, to begin finding ways to use the capabilities of the ARPANET in operational areas of the U.S. Government, particularly in the Department of Defense, the sponsor (through the Defense Advanced Research Projects Agency) of the Network.

One class of potential users of the Network has a certain set of problems which may not be faced by other users or potential users. The community of international security analysts - individuals who evaluate information about the behavior and capabilities of foreign nations - has a real need for the Network; a need which is not currently being filled. This paper attempts to investigate ways by which that need can be filled.

The report is prepared with the support of ARPA Contract No. DAHC15-71-C-0201, Dr. Robert A. Young, Principal Investigator; Dr. Michael R. Leavitt, Project Director. In preparing this report we have had the gracious assistance of Mr. Jean Iseli of the MITRE Corporation, Mr. Michael A. Padlipsky of M.I.T., Ms. Barbara Noble of UCLA and all of those whose reports, notes and information in the Network Information Center have proven invaluable.

SUMMARY

The thesis of this paper is that although the ARPA Network of Computers is a most versatile tool for many different kinds of problem-solving, it does not yet have the capabilities of serving the analyst in the international security area. The analyst has a need for convenient, reliable, relatively easy to learn, flexible systems, and for a moderately wide variety of programs and data. The ARPANET now provides convenience and flexibility, but not sufficient reliability, ease of learning, or data.

The paper begins by considering who the analyst is - his background and technical training, his computational needs, and how he currently satisfies those needs. The analyst in the international security area is a well-trained person whose skills have proven useful in understanding and recommending action about international affairs. He is trained as a foreign affairs generalist, but not in the use of computational tools. Moreover, he is not certain of the need for those tools in his profession. He works under substantial time pressure with a limited research/analysis staff and rarely has the time to commence major computational tasks with new data collections to answer operational questions posed to him. He requires reliable systems - systems in which he has confidence, and systems which will work when called upon. He needs systems which he can learn to use himself, and which, therefore, will not change radically over relatively short periods of time.

Good desc. of our analysts

The analyst currently has available moderately large batch processing systems - occasionally he will have access to a time-sharing system. They will be as reliable and stable as he needs to perform some functions,

but they take too long to use effectively. His current systems are generally inconvenient and difficult to learn, and most important, they rarely have a stockpile of accessible data which can be used to help answer his questions.

analyst

The ARPANET can do much to ease the convenience problem by providing time-shared access to a very large number of computers. The flexibility of the Network in permitting the use of a much larger number of programs, particularly those at UCLA, M.I.T. and Stanford University, can be a great aid. UCLA has a very substantial number of different statistical and modeling programs; M.I.T. has the CASCON system for local conflict, and the Consistent System for information processing in the most convenient form; Stanford has an analysis package linked to the Associated Press wire service.

However, the ARPANET is not yet sufficiently reliable for operational problems. Unreliability can come both from the Network itself, through its TIP's, IMP's, and channels; as well as from the individual computers on the Network. Although reliability has improved in both areas over the last few months, the analyst's need to be as certain as possible that a system will be available cannot yet be fulfilled. Nor is the Network particularly easy to learn to use. Connection from a local host computer obviates most communications problems while connection directly from a terminal may be difficult. The protocols for common user programs at various sites exist, but not in sufficient number to permit an analyst to query the "files" as he would with manual means. If answering a particular problem requires a major data collection and installation effort, as it would on most Network computers, the analyst will not use it to answer those questions, and will rely on his traditional sources.

Finally, the Network can be expected to develop its capabilities in all of these areas over the near future, thus increasing its utility to the analyst as well as the probability of its use. The two areas in which difficulties can be expected to remain are the reliability of local host computers, particularly those with primarily academic responsibilities, and the availability of data for the analyst's use. The paper concludes with a summary of literature about the Network.

I. WHO IS THE ANALYST?

The purpose of the paper is to assess the ARPANET's utility for the analyst in the international security community, so it is first necessary to understand some of the characteristics of the user. The information in this section is based on observations made during joint work with many analysts over the past few years as well as on various descriptions of the analyst and his functions. The section will proceed by first examining the analyst's background and technical training and will show that unlike individuals in other fields (systems analysis, operations research, economic policy, military strategy, etc.), the international affairs analyst has relatively little technical training - training in the use of analytical tools - but an increasing need for such tools and training in them. What the analyst needs to perform his job adequately will then be considered as far as computational support is concerned. In many areas, the analyst works in a very small group, in a crisis atmosphere and has little time for detailed research and learning about how particular systems work. His current computing situation will be discussed and its satisfactory as well as unsatisfactory aspects elaborated.

A. Analyst's Background

The International Security Analyst is a well-trained person whose skills have proven useful in understanding and recommending action about international affairs. These analysts have been properly considered to be some of the most valuable and competent individuals in the government for their ability to read, understand, synthesize and advise. Their role in government is now changing and, partially because of their past successes, analysts have not changed to fit the role.

Specifically, today's analyst is trained as a foreign affairs generalist with at least one advanced degree in some area of international affairs, whether he is a civilian, a military officer or a foreign service officer. The changes in methods of analyzing international affairs that have taken place over the last five to fifteen years have, in many cases, been too recent for the analyst to make direct use of. He is ambivalent about their utility for his operational responsibilities and, being well-trained in the substance of his field, is often hesitant to spend the time required to learn the newer techniques, which include the use of computers for statistical analysis, data display, modeling, and simulation.

It is most important to note that a major reason that the analyst is unwilling to spend great amounts of time on these techniques is that they have not had unqualified success in his field. The techniques are still controversial within the intellectual community of international affairs to which he belongs, and because of time pressures and career possibilities, an excursion into the realm of the controversial is not often rewarding. His ambivalence is not a matter of unintelligence or obstinacy, but rather, a realistic appraisal of where his valuable intellectual resources should be spent. His position is thus that proponents of new ways to make his work more efficient and more productive should come to him, rather than he to them.

B. Analyst's Needs

The analyst in the international security area is called upon to study a particular situation and produce recommendations for action by the U.S. Government. Although not usually in a crisis atmosphere, the analyst invariably works under time pressures, with little opportunity or incentive to "try something new" to see if it works. Further, the analyst rarely has a very large staff working for him to aid in his work.

He is expected to have the general store of information he needs in his own mind and his files, and to know where to go to get whatever else he needs. The analyst thus needs reliable sources of information - quickly and surely available - with an ease of access appropriate to a small staff. Any innovation that would attempt to serve that analyst will need to fulfill those particular requirements at a minimum.

The reliability of information sources has two components: the analyst must have confidence in them, and they must be available when needed. Confidence requires either that the analyst must have used them before or that someone must have demonstrated convincingly that they would have been at least as useful as personal prior use. Currently the analyst uses reliable sources, in that his colleagues use the same data and data-acquisition methods. The "files" which contain past actions of his agency, prior recommendations by him or his predecessors, and the safety of tradition are the primary sources of the analyst's information. To depart from these sources, the analyst needs to be convinced that they merit both his confidence and the confidence of his superiors.

But no source will be reliable if it is not available when needed. Time pressures on the order of hours or days prohibit the analyst from relying on information that can be acquired in a week, or which is currently available but he cannot access it for technical reasons. The current situation permits him or his assistant to go "into the files" for information and to have reasonable confidence that it is there. A system which is unlikely to yield that information will be distrusted and not used. A computerized information system might not be available to him during his normal working hours. Access to the system might require that he contact two or three technical personnel whose responsibilities include many other agencies or offices. Access to the information might require that he or his assistant learn how to interpret written information that was designed to be read by the statistician or computer programmer. Access to information might require that the analyst be transported to a site more than an hour's distance from his office. If any of these conditions exist, the information, no matter how believable, is unlikely to be used for operational purposes.

One additional characteristic of the source of the analyst's information needs to be mentioned. The procedures for access must be relatively stable. Once the analyst has made the investment of time to "learn" a method of information access, he must be able to rely on using that method for a reasonable period with no substantial new learning necessary. In other words, he needs an operational information and analysis system to help in his operational responsibilities.

C. Current Situation

The above sections outlined who the analyst is and what he needs. It is now necessary to specify what he has available to fulfill those needs in one specific area: computing resources for data analysis and presentation, since this is the major area in which the ARPA Network of Computers can be of assistance. A brief summary of the kinds of computing facilities generally available will precede an evaluation of how those facilities fulfill his need for reliable, stable, timely, and easy to learn information systems which provide the tools he needs to get his job done.

The most common facility currently available to the analyst is a large-scale computer, operating in a "batch processing" environment through a widely used "operating system." A large-scale computer might be considered to be a machine the size of an IBM 360/65 or bigger (IBM 370/155, UNIVAC 1108, Control Data 6000 series or Cyber 70 series, Honeywell 600 or 6000 would all be acceptable). Some users might have smaller machines with fewer computational [illegible] allow; others may have access to a time-sharing computer which certainly eases some access problems. The batch-processing system means that the analyst needs to have a particular piece of analysis prepared for him, submitted to the computer, retrieved, corrected, resubmitted and finally used. This process can take from one-half day to one week to be completed for many heavily used systems. The fact that the analyst has access to a widely used operating system means that he will be able to use several kinds of analysis software - programs already written for general statistical analysis or report preparation. This means that the analyst need not be a programmer, nor does he need a programmer on his staff.

The system fulfills two of his most important needs: stability and reliability. A large, widely used computer system will not change very much as far as the analyst is concerned. There may certainly be changes to the operating system, but they will rarely have a noticeable effect on the user's operation (except to lower or raise the cost of the jobs he does). Periodically, his analysis programs will be changed, but most often, the changes will be compatible with his previous procedures, by adding new procedures or by simplifying old ones. It is thus a stable system, and, once learned can be used for a period of time without relearning.

The system is reliable, for the most part, for two reasons. First, he is likely to have confidence in it. Any data that he uses is likely to have been collected and put on the system by his own office. Any programs that he uses are likely to have been used before by enough individuals in similar positions to minimize the risk of error. It is also reliable because it is going to be available when he needs it, or at least he will know if it is not able to be of use for a particular problem. A task which has a deadline of two weeks and will use precollected data may be perfectly acceptable for the use of even the most overloaded, the slowest or the smallest systems. One that requires results in two days will not even be attempted. In other words, the analyst knows the limits of his tools.

The system will be unsatisfactory, however, in its ability to provide convenience, ease of learning and sufficient data for handling new problem areas. The convenience aspect holds more for the user of batch systems than for time-sharing. A user with an immediate problem will not appreciate waiting two to four hours for output. It will be even more unsatisfactory to the extent that there were errors in the input which required resubmission. And it will be totally unsatisfactory if the analysis is sequential - if each stage of analysis requires the output from a previous stage. The situation is further degraded if the computer is some distance from the analyst's office or if the analyst must communicate his needs through an intermediary, be he an assistant, an in-shop consultant or worse, an outside consultant.

The time-sharing user has the advantage of much more rapid return on his efforts thus facilitating sequential analysis, ameliorating the affect of input errors and, usually, shortening the distance to his interface with the computer. For this user, the current situation is likely to be satisfactory as far as convenience is concerned.

Both the batch and the time-sharing user have the problem of training. Although there are an increasing number of exceptions, the effective use of current computers requires substantial training. One must understand what a computer can do, what its limitations are and what kinds of questions one can ask of it just to begin. Many analysts do not have this level of understanding, without which it is very difficult even to use a competent assistant. To make the best use of an analysis program, one must have spent some time learning, preferably from a person rather than a manual, just what that program can do, what kinds of input it requires, and how to interpret its output. If this learning takes place during the fulfillment of a particular task, particularly if there is a deadline approaching, the experience is certain to be frustrating at best.

Most computer systems now in use have little regard for the relatively computer-naive analyst. Few efforts are made to equip the programs with flexible input checking to forgive errors. Thus a learning process must take place. This process is not the defensible one of learning how to do a particular kind of analysis, but rather how to coerce a particularly ignorant tool to do that analysis for you. Most analysts have neither the time, the inclination or the perceived need to undergo that learning.

Finally, most current computers do not have libraries of data comparable to the "files." The "files" are those sources of information that the analyst has relied on for the information on which analysis is based. They may contain numbers, previous reports, intelligence information, or the results of prior analysis. To accomplish a new task, an analyst can generally rely on "the files" for the bulk of his data, leaving perhaps a small amount of information yet to be collected. The "files" are simply not available on any computers in the international security community (with some exceptions), so the analyst must place any needed information on the machine himself. After a period of time, the analyst may build up his own computerized "files," but their initial lack provides a barrier to current use. The process of building reliable information is difficult in any situation; on computers, however, the difficulty is compounded by the need for precision at every stage of the process.

As a result of the current computing situation, then, international security analysts are partially satisfied but generally not encouraged to use some of the newer tools in their operational problems. The ARPA Network offers much potential for resolving these problems leading to greater use of computer related analysis methods. The following section details how the ARPANET may be useful and where it too is unsatisfactory.

II. WHAT CAN THE ARPANET DO FOR THE ANALYST?

The thesis of this section is that the convenience and flexibility of the ARPA Network can be of some use to the international security analyst today, but that it is not yet reliable enough, sufficiently easy to learn and to use nor well enough stocked with data to permit its adoption. Although the emphasis of this section is negative on balance, the final section will illustrate that the current Network situation is still changing, and that the expectation is that the future of the Network is real potential for the analyst.

A. Convenience

The major purpose of the ARPA Network of computers is convenience and flexibility:

> The goal of the computer network is for each computer to make every local resource available to any computer in the net in such a way that any program available to local users can be used remotely without degradation. That is, any program should be able to call on the resources of other computers much as it would call a subroutine. The resources which can be shared this way include software and data, as well as hardware.[1]

On the whole, the Network has succeeded in its goal. It is convenient in two distinct ways. If a user has local access to a "host computer," he

---

[1] Larry G. Roberts and Barry Wessler, "Computer Network Development to Achieve Resource Sharing," in AFIPS Conference Proceedings, Vol. 36, 1970 Spring Joint Computer Conference (Montvale, N.J.: AFIPS Press, 1970), p. 543.

Approved For Release 2004/09/23 : CIA-RDP79M00096A000500010013-4

can [illegible]ve. The access is achieved by a network of high-speed communication channels which links over twenty-five different computers throughout the United States, Hawaii, and most recently England and Norway. By an appropriate use of the local computer's "file transfer protocol" program, a user can transfer large amounts of data from any other computer on the Network to his own for processing. Thus, if an analyst at the RAND corporation wished to perform a statistical analysis at his own site using data normally available only on the MIT-Multics computer, it would be a relatively simple matter to have the data brought over specifically for use by his program.

The more striking convenience comes for the user who does not have direc access to his own computer, however. This user may use a standard telephone line to call a Terminal Interface Message Processor (TIP) whic can then connect to <u>any</u> computer on the Network. Some TIP's are for the sole use of their owners (for example, TIP's at the USN Fleet Numerical Weather Central in Monterey, or at the USAF Environmental Technical Ap plication Center at the Navy Yard Annex in Washington). Others function (currently) as a more general resource (for example, TIP's at MITRE in McLean, Virginia and at the National Bureau of Standards in Gaithersburg, Maryland). The analyst with access to a TIP from a communications terminal (such as a Teletype, an IBM 2741, a General Electric Terminet or a Hazeltine 2000) can "look like" a local time-sharing user of any "server" computer on the Network. As must other local users, he must make arrangements for access with the local computer personnel, but this is facilitated by the presence of a Network Liaison person at each site. In addition, there are several servers which provide experimental usage of their systems at no direct charge to the user (these include MIT-Multics, Stanford's Artificial Intelligence Laboratory, and the Network Information Center at Stanford Research Institute).

If the programs and data that an analyst needs for an operational task are on the Network, the convenience of using that information, either at the location of the information, or at the analyst's home site (if a member of the Network) cannot be surpassed.

B. Flexibility

Flexibility in this context means the ability to perform or to help in the performance of a wide variety of computational tasks. The Network gives the analyst more different computer programs and ways of using those programs than he will be able to use. (The Network is so flexible in this sense, that it almost becomes a disadvantage, as discussed in the next section.)

Of the over twenty-five computers currently on the Network designated as "servers," a few provide the bulk of the readily available programs that an analyst is likely to need. These are the Campus Computing Network at the University of California, Los Angeles, the Artificial Intelligence Laboratory at Stanford University and the Multics computer, a part of the Information Processing Center at the Massachusetts Institute of Technology.[2]

1. Campus Computing Network at UCLA

The UCLA Campus Computing Network (UCLA-CCN) is served by the largest and currently most powerful computer on the Network, an IBM 360 model 91. The size of the computer permits readily available

---

[2] Many other computers have many other resources of general interest; however, for the current and projected future needs of international security affairs analysts, these three are of the greatest interest.

storage of some of the most commonly used statistical and modeling programs available. Its statistical packages include the set of Biomedical Computer Programs (BMD) which, in spite of its name, is perhaps the most commonly used set of programs for nearly all kinds of statistical data analysis. BMD is not the easiest program set to use or interpret, but it is likely that if an analyst is familiar with only one set of programs, it will be BMD. UCLA-CCN also has the Statistical Package for the Social Sciences (SPSS) available. SPSS is substantially easier to learn and to interpret because it is a single integrated package. DATA-TEXT, (a package developed at Harvard) and the Statistical Subroutine Package (SSP, an IBM product) are also available.

UCLA-CCN also has the two most common computer simulation programs available. SIMSCRIPT II.5 (a proprietary simulation language developed by C.A.C.I.) and GPSS (an IBM product) can be used by the experienced user for a very wide variety of tasks.

UCLA-CCN permits two modes of operation: time-sharing and remote batch processing. Both modes can be done from a computer terminal, although they have very different characteristics. Time-sharing is performed under the Time-Sharing Option (TSO), with which the user may write programs in FORTRAN, PL/1, and BASIC. The time-sharing user may create and edit files of data or textual information, and use these as input and to store output from programs. Remote-batch processing is done from the terminal, but in this mode of operation, the user creates a file of information which the computer processes independently of user interaction. It is as if the user created a batch processing job and submitted it to the computer. As with batch jobs generally, the user must wait for the computer to finish processing before he is able to examine the output. The system is faster than normal batch processing, but the lack of direct user intervention can be a problem. Most of UCLA-CCN's progr... ed at the beginning of this subsection.

2. Stanford University - Artificial Intelligence Laboratory

At first glance, there is no particular reason why an artificial intelligence operation should be of use to an international security affairs analyst. However, one of their projects involves making the Associated Press Wire Service available to users of their system. The AP wire is available in two ways. Through a program named "HOT," the user may have his terminal behave like an AP terminal, with national and international news stories printed out at the same time that they appear on terminals in news rooms across the country. This is more a form of entertainment than a useful adjunct to analysis. The useful product is a program entitled "APE" (Associated Press Extractor) which retrieves information from a constantly updated one-day file of AP stories. Information can be retrieved by topic ("SALT TALKS," "MIRV," "ABM"), by country or city ("USSR," "MOSCOW," "WEST BERLIN") by individual ("NIXON," "MAO," "SADAT") or by any combination of words that may have appeared in a news story. Once retrieved, the information can be displayed at the terminal or stored in a file. Since only one day's worth of stories is available at any time, it is necessary to perform the analysis daily, or take advantage of the "standing order" facility by which stories of interest are automatically stored in a user's file. Files thus created can be read from the computer, or sent, via the File Transfer Protocol program, to another computer for content analysis or long-term storage.

Stanford Artificial Intelligence Laboratory (SAIL) is one of the systems which provides free experimental usage, thus facilitating investigation prior to commitment of resources. Since it is located in the Pacific Time Zone, it is relatively unused during the morning (on the East Coast) and thus an attractive system. The manual for using the AP system is well-written and brief.[3]

3. Multics at the Massachusetts Institute of Technology

MIT has four different computers currently on the Network, Multics being the most often used as well as the most powerful. Multics began as an experimental system to investigate the idea of a "computer utility" - a system which could simultaneously support, at different levels of service and different pricing arrangements, a large number of users. Multics has recently been moved from an experimental computer (a Honeywell Information Systems 645) to a production version which is being commercially marketed (HIS 6180). It is of great interest for its means of protecting the files and security of its users and its effective "unlimited" storage capacity (from the user's perspective), but for the international security analyst, two particular programs are of interest: CASCON and the Consistent System.

CASCON is a single program which provides systematic information about 52 local conflicts since the second World War, to the user at the terminal. The user can enter information about new conflicts to determine how comparable they are to the older ones (the "file"). Basically an information retrieval system, CASCON[4] was developed for the U.S. Arms Control and Disarmament Agency, and has not been used to any great extent

---

[3] See Martin Frost, "Reading the Associated Press News" (Stanford, California: Stanford Artificial Intelligence Laboratory, July 1973).

[4] Developed by Lincoln Bloomfield and his associates at MIT's Center for International Studies. See Lincoln Bloomfield, Robert R. Beattie, and G. Allen Moulton, "CASCON II, Computer Aided-System for Handling Information on Local Conflicts: User's Manual" (Cambridge, Mass.: Center for International Studies, September 1972).

because of its need for a particular computer system for its operation (although it has been converted for use on the University of Michigan system, not a member of the ARPANET). CASCON may prove of interest to some members of the international security community.

The Consistent System, on the other hand, is a large set of programs designed to help behavioral scientists use computers in their research. A part of the ARPA-funded Cambridge Project, the Consistent System is available only on Multics. Of particular interest for data analysis and presentation is the JANUS subsystem of the Consistent System. JANUS permits a limited number of kinds of statistical analysis, but is one of the most flexible, forgiving programs currently available. Other functions that the Consistent System is able to fulfill include programs for computer content analysis of documentary material and, for the more sophisticated user, a very complete system for classifying and manipulating stored information.

Many other kinds of programs are available on the Network, but those listed above may properly be considered as the most relevant to the needs of most international security analysts. It is fair to say that they represent a larger base of computing power than is currently available on the large operational computers in the various agencies.

One important inflexibility remains to be noted that will prevent the ARPANET from replacing currently used computing resources for some tasks. The ARPANET is, of course, an insecure network. No classified analysis can be done directly through the Network nor can it be done on the computers in the Network, unless the analysis is masked in such a way as to prevent any possibility of determining the content of the analysis. For these purposes, analysts will need to continue to use secure installations in the manner prescribed by the appropriate regulations.

C. Reliability

The ARPANET is not yet reliable enough to serve as an operational tool. It is unreliable both in its communications system, and in its components, the host computers. This strong statement admits to some exceptions, but is based on the rather unusual needs of the international security analyst. As elaborated above, an analyst can be called upon for a report which is due a matter of days or hours from the time of its request. For any computer system to be of use in this environment, it must be able to guarantee its own reliability when called upon. The ARPANET cannot yet do that for two reasons: unreliability of the Network communications and unreliability of the hosts.

1. Network Unreliability

The Network is a complex linking of computers through intermediate computers through which data are transmitted. The TIP discussed above is a generalized version of the Interface Message Processor (IMP) a computer which receives information from the various host computers to be passed on to other IMP's and then to their proper destination. Given the number of stages through which a particular set of data must pass to reach its destination, the ARPANET is reliable for most purposes.[5] However, the Network can become unavailable to the user for a variety of reasons, including, malfunction of his own terminal, his host computer, or the IMP at his host. The first two problems cannot properly be attributed to the

---

[5] See Howard Frank, Robert E. Kahn, and Leonard Kleinrock, "Computer Communication Network Design - Experience with Theory and Practice," in AFIPS Conference Proceedings, Vol. 40, 1972 Spring Joint Computer Conference (Montvale, N.J.: AFIPS Press, 1972), pp. 255-270 for a more complete discussion of the reliability problems in the ARPANET.

Network, although the third can. Further, because of the way the Network is designed, if an IMP through which messages must pass to reach their destination malfunctions, the whole Network may be down as far as some users are concerned (although the topology of the Network was designed to minimize that possibility). If an analyst is communicating with a TIP, he avoids problems of a local host not being available, at the cost of introducing a new set of problems. TIP's are functionally equivalent to IMP's for Network communications, and since they are more complex than IMP's (being able to communicate directly with users) they are more likely to suffer failure.

Thus, the user has a small, but not insignificant likelihood of not being able to use the Network communications system when it is needed to perform a particular task. This certainly will not occur very often (statistics, of course, are unavailable given the lack of users in this position) but one or two such experiences may be sufficient to discourage further reliance on the Network. Coupled with the potential for remote servers to be unavailable, reliability is a problem.

2. Remote Server Unreliability

The various servers on the Network differ greatly in their individual reliabilities. Most, however, share a common characteristic that as academic computing centers, they are not generally required to maintain the same standards of availability as commercial or military computers. In general, users of academic computing centers have fewer time pressures than commercial and military users, while academic computer center management has fewer incentives to maintain availability than commercial management.

Academic research and classroom use is very rarely dependent on the computer's availability at any given moment. Further, academic users do not generally have the option of going to a competing service if their computer is not available as often as they wish. Commercial users generally do not require full availability, but they have the competitive bargaining position. Military users often require availability, and can enforce such through the structure of their organization.

The point of the above is that there is no particular reason to expect that academic hosts on the ARPANET will be available as often as might be needed by the international security affairs analyst. This expectation is born out by experience. However, unreliability, if it were the only problem, might not be sufficient to disqualify the Network from being an operational tool for the analyst. Indeed, if this were the only problem, a very good case could be made for the Network as a standard research tool, but one that simply could not be relied upon in urgent situations as is the case with current computers. The remaining problems are such that reliability becomes only one item to be considered.

D. Ease of Use

Like reliability, ease of use is a function both of the Network communications and the individual hosts. The ARPANET is difficult to learn to use creatively, although performing the same task repeatedly is relatively easy. The rewards of learning to use the Network as a whole are potentially great, but the initial learning is difficult enough to pose a problem to the international security affairs analyst.

1. Using the Network

The two modes of using the Network - via the TIP from a terminal, or via an IMP through a host computer - are different experiences for the user. The TIP is a very simple machine - strict and unforgiving with a preference for numerical identification of servers rather than their names. Getting access to a host computer through a TIP may be as simple as a one line command:

@LOGIN 11

which tells the TIP to open a connection to SAIL (44 for MIT-Multics and 65 for UCLA-CCN), after which the user is effectively talking through the TIP to the computer he wishes to work with. Or, depending on the kind of terminal the analyst is using, it may require the following dialogue:

@DEVICE CODE EXTRAPADDING

@TRANSMIT ON LINEFEED

@ECHO LOCAL

@LOGIN 44

which tells the TIP that MIT-Multics is to be logged into; that Multics prefers to receive large amounts of information at a time; that the terminal has a need for delays before carriage returns are printed; and that the TIP should echo everything typed into it. To repeat, the analyst who is only using one computer can learn the proper sequence of commands to the TIP. The user who wishes to access more than two or three computers at different times may have difficulty remembering which machines require which commands.[6]

The IMP user has a much more pleasant experience. The local host

---

[6] See the "Terminal Interface Message Processor User's Guide" (Bolt Berenek and Newman Report 2183, June 1973) for complete details.

through which the analyst would communicate with the Net has a program called TELNET which performs the same function as the TIP. However, TELNET, since it is based on a larger computer, can be much more flexible. Very often it may only be necessary to tell the local computer to:

#connect to multics

for the computer to know exactly what details are necessary and do them. The next message to the user would be from Multics.

Learning to use the Network Communications, then, is relatively difficult for an analyst with little interest in communications for its own sake who may need to use several different computers. If the analyst needs only one remote computer, or if the analyst is going through his local host, or, finally, if the analyst can have some member of his staff make the required connection, communication with the Network should not be too difficult.

2. Using the Servers

The ease of using different servers varies from the almost TIP-like rigidity of UCLA-CCN under remote batch, to character by character checking of the user's input on the series of TENEX computers (TENEX is an operating system available on six of the Network computers). If there were no common source of information on how to use the various Network computers, none but the most determined would learn more than one or two of them. Fortunately, there are two ways by which the relatively unsophisticated user (like the analyst) can be helped.

First is a set of conventions which are common to all of the above discussed network computers: the HELP command and the @SYNC operation.

UCLA-CCN, SU-AI and MIT Multics all accept the HELP command at very many stages in processing. When the user types HELP, or HELP XXXX (where XXXX is a topic like LOGIN, FORTRAN, etc.), the computer will reply with a paragraph informing the user either how to get more information, or how to do what he needed to do. This commonality - an ability to type HELP and get it - is one of the most valuable assets of the ARPANET.

On occasion, the user will get a large volume of output at his terminal and not know how to stop the machine from spewing out more. By typing "@SEND SYNC," the TIP is instructed to send the appropriate message telling the computer to stop doing what it is doing, and wait for a new request from the user. This too is a most useful common element. In addition to the above, the new user will be able to learn a great deal by signing onto the Network Information Center (NIC), an on-line storehouse of vital information.

The NIC is physically located at Stanford Research Institute, Augmentation Research Center (SRI-ARC) and employs a PDP-10 computer (under the TENEX operating system). It is entered by logging into NIC, SRI-ARC or (from the TIP) number 2. One must first obtain an account, but once this is accomplished (contact ARPA for such information), the user may immediately access a set of data files collectively entitled "The Resource Notebook" (also available off-line). The Resource Notebook contains an up-to-date set of basic information on how to use each computer on the Network. The Resource Notebook can be queried about computers, programs, its own documentation and other topics.

However, the analyst who needs to solve a problem will be marginally interested in using a wide variety of computers. He will need complete documentation on the particular program he wishes to use. He must rely on the local computer center for this information, particularly if it is a locally produced product (like the Consistent System on Multics or the Associated Press system at Stanford). The quality of this documentation will vary substantially, from the excellent documentation of the Consistent System and Associated Press, to the inadequate documentation for many programs at UCLA-CCN.

As a result of the above situation, the analyst with limited goals can use the network and its computers quite adequately at present. Where the number of tasks increase, as one might expect given the resources on the network, and as the number of computers needed to perform those tasks increase, the analyst will find himself increasingly confused as to how each individual part of the overall Network system works.

E. Data Availability

In the previous section, the data needs of the analyst were stressed. It should not be surprising to note, very simply, that there is virtually no data of use to the international security affairs analyst on the Network with the sole exception of the data base for the CASCON program discussed above.[7] This should not be surprising because relatively few data collections were created at universities or agencies with computers on the Network. To the extent that the analyst needs prior data already available, he is quite simply not going to find such on the ARPANET.

---

[7] An additional data base will be available by the end of 1973 when the UCLA Center for Computer-based Behavioral Studies (CCBS) comes into the Network. Current plans are for the World Event/Interaction Survey, an ARPA-funded collection of information on international interactions among all nations of the world since 1 January 1966, to be available through CCBS.

III. WHAT WILL THE ARPANET BE LIKELY TO BE ABLE TO DO FOR THE ANALYST?

Since predictions are generally risky, and technological and policy predictions are among the riskiest, this section will be relatively brief. In each of the five areas of the previous section, the apparent direction of change will be stated, and, if that is not likely to be of use to the analyst, suggestions for what could be done will be briefly made.

A. Convenience

The ARPANET is very likely to become even more convenient for the analyst to use. Hosts are being added to the Network at a regular rate, and for the analyst based in Washington, the possibilities of direct access to the Network through a host will increase. Further, the kinds of terminals that will be able to communicate directly with the TIP will increase in number, thus providing access to those offices with equipment not currently compatible. Since this area is currently satisfactory, the outlook is particularly optimistic.

B. Flexibility

The Network can be expected to become increasingly flexible as more hosts acquire and develop programs of direct use to the analyst. As suggested above, too much flexibility may have its negative aspects in that the potential user will have difficulty assessing what is useful. However, further contracts such as that between ARPA-IPT and the MITRE Corporation, through which advice is given to potential users by individuals who know both the needs of the individuals and the capabilities of the Network, can be expected to ease the "problems" of flexibility.

C. Reliability

The reliability picture is mixed. Reliability of the communications system can be expected to improve to the point where the analyst will be able to take for granted the ability to access most servers upon demand. The reliability of the servers, however, can not be expected to increase nearly as much, since the problem is a result of different requirements on the part of the analyst and the users of the host computers. This is particularly true for academic centers where reliability, as has been discussed, is desirable but not crucial. Increasing the number of commercial servers on the Network would be likely to ameliorate this problem. Several companies are now or will soon be on the Network (including Bolt Berenek and Newman, the developers of the IMP, the TIP, and the TENEX operating system; and Tymshare Corp., a successful time-sharing vendor) and this can be expected to benefit the reliability of the system, and the expectation that the analyst will be able to use the resources on demand.

D. Ease of Learning and Use

This area is particularly difficult to forecast, since it depends upon policies which will be made over the near future. To the extent that emphasis is placed on ease of use - improving the TIP software to be somewhat simpler, and encouraging hosts to utilize more common practices in the operation of their most often used software, the Network will become easier to use. There is, in particular a need for greater Network documentation of user programs and data available at the various servers. Improvements are being made regularly in this area, but there is much to be done beyond the current practice of simply listing the program name and the contact person. The analyst who receives a list of more than six or seven similar programs with similar names or one-line descriptions would benefit substantially from having more complete documentation available through the Network (at the NIC) rather than being forced to contact the various individuals who are responsible for further information. Such individuals should become important after a particular program has been chosen, rather than at the time of deciding.

E. Data Availability

The ARPANET currently has plans to bring on-line a unique memory for one of its already operational hosts, the PDP-10 at the Computer Corporation of America (CCA-TENEX), also called the DATACOMPUTER. This memory will be in a read-only form (once it is "written" it will be permanent) and should be available to any Network user upon appropriate financial arrangements. The Network will have the ability to store vast amounts of data, some of which may be of use to the international security affairs analyst. As of now, however, there are no public plans for "stocking" the DATACOMPUTER or any other computer, for that matter, with data specifically needed by the analyst. This is a matter of some concern, in that even if all of the above criteria for use are met, without a useful "memory" the likelihood of an analyst adopting the Network for operational activities is low.

Although any future Network-available data is unlikely to be classified, analysts currently use large amounts of unclassified information in their work, including the population and production data sets compiled by the U.S. Departments of Agriculture and Commerce, information on military expenditures compiled by the U.S. Arms Control and Disarmament Agency and information on the status of military forces developed by the International Institute for Strategic Studies. It would be a relatively easy matter, technically, to install some of these data sets, or others on one of the Network servers while waiting for the operational status of the DATACOMPUTER.

As work at C.A.C.I. and elsewhere has demonstrated, there is a substantial number of data sets available for such installation. It may or may not be the proper function of ARPANET personnel and sponsors to support such activity, but, if the Network is to be used by the international security affairs analyst, such will be very important.

IV. ADDITIONAL INFORMATION

The ARPANET and its functions are well documented and a great amount of regularly updated information is available. This section details some of the sources for further information.

A. ARPANET Technical Information

The following papers were presented at the Spring Joint Computer Conference, 1970, and were published in the AFIPS Conference Proceedings, Volume 36 (Montvale, N.J.: AFIPS Press, 1970).

Carr, C. Stephen, Stephen D. Crocker and Vinton G. Cerf "HOST-HOST Communication Protocol in the ARPA Network " pp. 589-597.

Frank, H., I.T. Frisch and W. Chou "Topological Considerations in the Design of the ARPA Computer Network" pp. 581-587.

Heart, F.E., R.E. Kahn, S.M. Ornstein, W.R. Crowther and D.C. Walden "The Interface Message Processor for the ARPA Computer Network" pp. 551-567.

Kleinrock, Leonard "Analytic and Simulation Methods in Computer Network Design" pp. 569-579.

Roberts, Lawrence G. and Barry D. Wessler "Computer Network Development to Achieve Resource Sharing" pp. 543-549.

The following papers were presented at the Spring Joint Computer Conference, 1972, and were published in the AFIPS Conference Proceedings, Volume 40 (Montvale, N.J.: AFIPS Press, 1972), and as a separate booklet available through ARPA-IPT (Information Processing Techniques).

Crocker, Stephen D., John F. Heafner, Robert M. Metcalf and Jonathan B. Postel "Function-oriented Protocols for the ARPA Computer Network" pp. 271-280.

Frank, Howard, Robert E. Kahn and Leonard Kleinrock "Computer Communication Network Design - Experience with Theory and Practice" pp. 255-270.

Ornstein, S.M., F.E. Heart, W.R. Crowther, H.K. Rising, S.B. Russell and A. Michel "The Terminal IMP for the ARPA Computer Network" pp. 243-254.

Roberts, Lawrence G. "Extensions of Packet Communication Technology to a Hand Held Personal Terminal" pp. 295-298.

Thomas, Robert H. and D. Austin Henderson "McROSS - A Multi-computer Programming System" pp. 281-293.

Other sources of technical information are available, and are either noted in the references to the above papers, or are available through ARPA-IPT or the MITRE Corporation (Mr. Jean Iseli).

B. ARPANET Resources

The most important single source of information on the resources of the ARPANET is the "Resource Notebook" available through the Network Information Center, at the following location:

Ms. Jeanne North
Stanford Research Institute
Augmentation Research Center
333 Ravenswood Avenue
Menlo Park, California 24025

This document will provide basic information on all Network sites (host computers, servers, IMP's and TIP's). For further information about any particular site, the Network Liaisons for each should be contacted. For those sites discussed above, the following people are available:

Mr. Michael A. Padlipsky (Multics)
Information Processing Center
Building 39
M.I.T.
Cambridge, Mass. 02139

Mr. Robert Bell (UCLA-CCN)
Campus Computing Network
Math-Science Building
University of California
Los Angeles, California 90024

Mr. Ralph Gorin (SU-AI)
Stanford Artificial Intelligence Laboratory
Computer Science Department
Stanford University
Stanford, California 94305